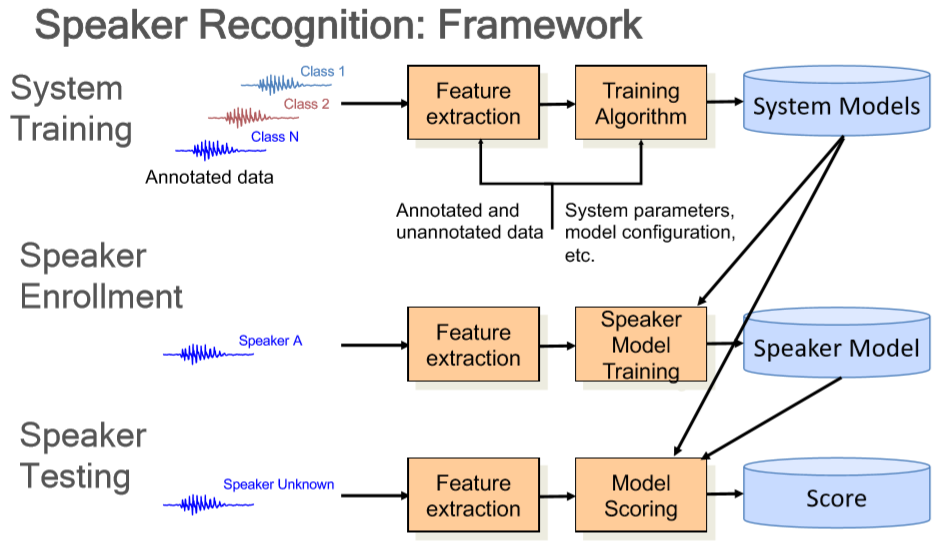
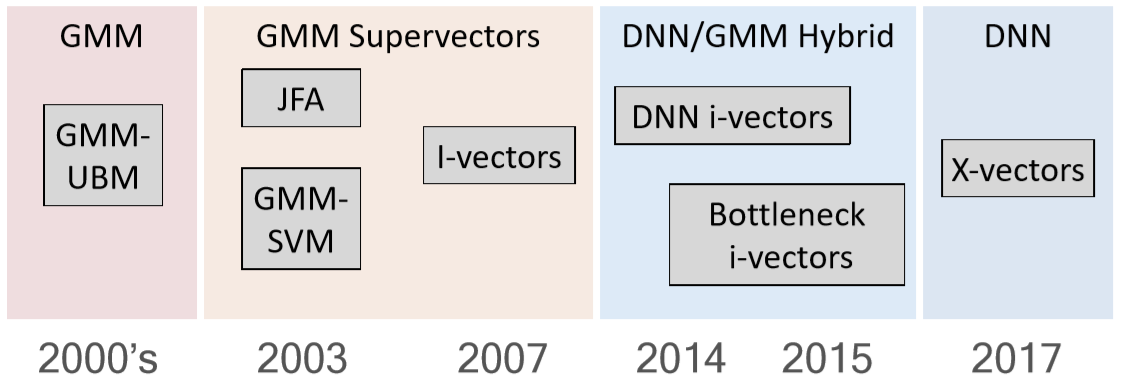
# 相关的问题

## 解决框架

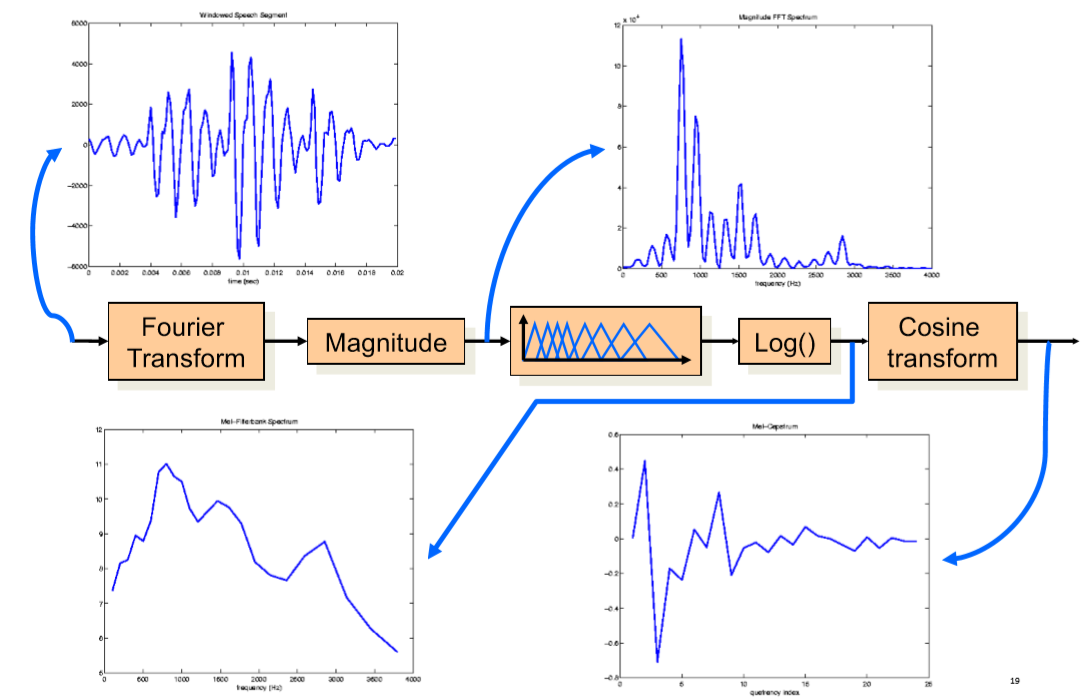


### 细节

## 方法的演进



## 音频数据的处理



* 傅里叶变换
* 放大
* 取对数log()
* 余弦变换

### 音频文件的存储细节

## 提取哪些特征

从低级到高级：

* 呼吸、鼻
* 韵律、节奏、语调、音量调节
* 语义、发音、特质

## 检测模型的时候如何排除内外因素的干扰？

* 外在
* 录音设备
* 声环境
* 传输的通道
* 内在
* 健康状况
* 紧张
* Role?

## 条件方面的挑战

* 校准-确保系统在部署后符合预期的错误率
* 远距离语音-低SNR音频和混响
* 鸡尾酒会问题-可以以较高精度识别一个人所讲的话，但是当说话的人数为两人或者多人时，语音识别率就会极大的降低

## 数据

VoxCeleb：

全自动的数据采集